

Study framing influences crowdsourced rates of problem gambling and alcohol use disorder

Angus, Douglas J; Pickering, Dylan; Keen, Brittany; Blaszczyński, Alex

Published in:
Psychology of Addictive Behaviors

DOI:
[10.1037/adb0000687](https://doi.org/10.1037/adb0000687)

Licence:
Other

[Link to output in Bond University research repository.](#)

Recommended citation(APA):

Angus, D. J., Pickering, D., Keen, B., & Blaszczyński, A. (2021). Study framing influences crowdsourced rates of problem gambling and alcohol use disorder. *Psychology of Addictive Behaviors*, 35(8), 914-920.
<https://doi.org/10.1037/adb0000687>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

For more information, or if you believe that this document breaches copyright, please contact the Bond University research repository coordinator.

Study Framing Influences Crowdsourced Rates of Problem Gambling and Alcohol Use Disorder

Douglas J Angus^{1,2}, Dylan Pickering², Brittany Keen², and Alex Blaszczynski²

¹School of Psychology, Bond University

²School of Psychology, The University of Sydney

Accepted for publication in Psychology of Addictive Behaviors.

© 2021, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication, via its DOI: 10.1037/adb0000687

The preregistered protocol for this study is available via <https://osf.io/2wafm>. Anonymized data for this study are available via <https://osf.io/d7h6e/>.

Correspondence concerning this article should be addressed to Douglas J Angus, School of Psychology, Bond University, Gold Coast 4227, Australia

Contact: dangus@bond.edu.au

Abstract

Objective: Crowdsourcing is an increasingly popular source of participants in studies of problem gambling. Studies with crowdsourced samples have reported prevalence rates of problem gambling between 10 to 50 times higher than traditional sources of estimates. These elevated rates may be due to study framing motivating self-selection. In this preregistered study, we examined whether study framing influences self-reported problem gambling severity and harmful alcohol use in a sample of participants recruited from a popular crowdsourcing website.

Method: Two recruitment notices for an online questionnaire were placed on Amazon Mechanical Turk (MTurk). Notices were framed as “Gambling and Health” or “Alcohol and Health”. Only participants who passed data checks were retained for confirmatory analyses ($N = 564$; 44% of recruited participants). Participants in the gambling framing ($N = 261$) and in the alcohol framing ($N = 303$) were compared on the Problem Gambling Severity Index (PGSI) and Alcohol Use Disorders Identification Test (AUDIT).

Results: Problem gambling rates and severity scores were significantly greater for participants in the gambling framing compared to those in the alcohol framing. Self-reported scores of harmful alcohol use were significantly greater for participants in the alcohol framing compared to those in the gambling framing, but there was no significant difference in prevalence rates for harmful alcohol use.

Conclusions: Study framing is an important consideration for gambling and alcohol research. We found that study framing may substantially increase the observed rates of problem gambling severity in crowdsourced samples, potentially via encouragement of self-selection.

Keywords: Crowdsourcing, problem gambling, addiction, alcohol

Public Health Significance Statement

This study highlights the importance of appropriate recruitment advertisements when using crowdsourced participants. The results suggest that using advertisements which mention gambling or alcohol may substantially increase the reported severity of disorders associated with these behaviors or substances. These increases may be due to self-selection or cognitive biases, and studies using crowdsourced participants should consider this effect when interpreting their results.

Psychological research has traditionally relied on convenience samples, typically first year university students in Western democratic countries (Henrich et al., 2010a). Population-based surveys and questionnaire studies targeting defined subpopulations are often expensive to conduct, given costs involved with random computer-assisted telephone and face-to-face interviews. Consequently, there has been an increase in the use of crowdsourcing websites such as the Amazon Mechanical Turk (MTurk) to recruit large samples for psychological research (Paolacci & Chandler, 2014). There is also growing concern about how representative these samples are of the general public, with respect to basic phenomena (Henrich et al., 2010a, 2010b) and specialist psychological phenomena and behavior (Gainsbury, Russell, & Blaszczynski, 2014).

Although several recent studies have suggested crowdsourcing can yield high quality data similar to that from traditional recruitment streams (Casler et al., 2013; Kim & Hodgins, 2017; Mishra & Carleton, 2017) others have yielded high rates of problem gambling and harmful alcohol use. The rates of self-reported problem gambling in past-year gamblers sampled from sources such as MTurk have ranged from 27.5% (Poole et al., 2017) to 51.7% (Hollingshead et al., 2016) compared to prevalence study rates ranging from 1.0% of past year gamblers (Gainsbury, Russell, Hing, et al., 2014; Wardle et al., 2011) to 2.1% of monthly gamblers (Armstrong & Carroll, 2017). The rates of harmful alcohol use in crowdsourced samples also appear to be elevated compared to traditionally sampled prevalence studies. These diverging estimates have led to concerns about the validity, reliability, and utility of using crowdsourced samples (Dennis et al., 2018; Dreyfuss, 2018; Hays et al., 2015; Kennedy et al., 2016; Yeager et al., 2011)¹.

¹ Note that we are not arguing that crowdsourcing should never be used, or that it is inappropriate. Rather, that the rates of psychopathology observed are substantially higher than reported elsewhere. The underlying cause for these elevated rates is unclear, and to the best of our knowledge, as yet unexplained.

One potential source of variation in crowdsourced rates may be the framing of study advertisements. In a study using door-to-door, face-to-face, and CATI surveys, rates of problem gambling were more than twice as high when the survey was framed as being about gambling compared to it being about “health and recreation” (Williams & Volberg, 2009). These elevated rates represent an important source of bias. Though, it is not clear if increased rates result from individuals interested or personally invested in gambling or alcohol (e.g., self-selection), or due to demand characteristic biases resulting from cognitive framing effects (Moss & Ranson, 2006; Norton et al., 2009; Tversky & Kahneman, 1981; Weinstein et al., 2018). The potential impact of study advertisement is a necessary and important consideration for crowdsourced studies.

In the present study, we aim to answer a simple question: to what extent does stating that a study is about a potentially addictive behavior or substance lead to greater reported rates of the associated disorder in a crowdsourced sample? Specifically, we examined the extent to which instructions describing the nature of the study altered rates of problem gambling and harmful alcohol use in crowdsourced participants. Our objective was to compare responses to the same questionnaires under conditions where the study was framed as investigating gambling behaviors as opposed to alcohol consumption. We tested the following hypotheses²:

H1a. The severity of self-reported problem gambling will be greater in a study framed as being about gambling compared to a study framed as being about alcohol.

H1b. Rates of problem gambling will be greater in the gambling-framed study compared to the alcohol-framed study.

² Additional hypotheses were preregistered for other variables, but following the reviewer comments, these have been removed from the main body of the manuscript. A complete, transparent reporting of these hypothesis and variables is available in the supplementary material.

H2a. The severity of self-reported alcohol use disorder will be greater in a study framed as being about alcohol compared to a study framed as being about gambling.

H2b. Similarly, rates of harmful alcohol use will be greater in the alcohol-framed study compared to the gambling-framed study.

The specific hypotheses, study protocol, and planned analytical choices were preregistered (<https://osf.io/2wafm>). Note that the wording of hypotheses have been altered for clarity and consistency.

Method

Participants

The initial sample comprised 1,173 participants accessing an online Qualtrics survey recruited through Amazon MTurk. Participants were only able to take part if they were aged 18 years or older, and residents in the following English-speaking countries: United States, Canada, United Kingdom, Ireland, Australia, and New Zealand. Participants received 50 cents credited to their Amazon account in exchange for survey completion. The study was approved by the University of Sydney Human Research Ethics Committee, and all retained participants provided digital consent.

One hundred and thirty-two individuals opened the Qualtrics survey but were excluded for either not consenting to, or not completing the survey. An additional 31 participants were excluded for not providing complete responses to key variables, yielding a maximum analyzable sample of 1,010. A further 446 participants were excluded for failing response-integrity and/or attention checks, resulting in an analytic sample of 564.

Demographic details for the analytic sample are presented in Table 1. Demographic details for the full sample and their response-integrity and attention check data are available in the supplementary information.

Table 1. Demographic characteristics for Gambling condition (N = 261) and Alcohol condition (N = 303) participants.

	Gambling condition N (%)	Alcohol condition N(%)
Age [mean (SD)]	38.03 (12.48)	37.95 (12.01)
Gender		
Male	128 (49.04)	120 (39.60)
Female	133 (50.96)	181 (59.74)
Other (please specify)	0 (0.00)	2 (0.66)
Marital Status		
Married	107 (41.00)	115 (37.95)
Living with partner/de facto	34 (13.03)	44 (14.52)
Widowed	5 (1.92)	3 (0.99)
Divorced or separated	19 (7.28)	30 (9.90)
Never married	96 (36.78)	111 (36.63)
Education		
Year 10 or equivalent	1 (0.38)	2 (0.66)
Year 12 or equivalent	50 (19.16)	71 (23.43)
A trade/technical certificate or diploma	39 (14.94)	37 (12.21)
University or college degree	125 (47.89)	134 (44.22)
Postgraduate university qualifications	46 (17.62)	59 (19.47)
Employment Status		
Full-time	161 (61.69)	162 (53.47)
Part-time	27 (10.34)	45 (14.85)
Unemployed	16 (6.13)	28 (9.24)
Full-time student	10 (3.83)	17 (5.61)
Self-employed	27 (10.34)	30 (9.90)
Disability or other (not aged) pension	4 (1.53)	5 (1.65)
Retired (including aged pension)	11 (4.21)	8 (2.64)
Other (please specify)	5 (1.92)	8 (2.64)
Occupation		
Manager	37 (14.18)	37 (12.21)
Professional	100 (38.31)	84 (27.72)
Technician or trade worker	10 (3.83)	21 (6.93)
Community or personal service worker	9 (3.45)	15 (4.95)
Clerical or administrative	25 (9.58)	41 (13.53)
Sales	23 (8.81)	20 (6.60)
Machinery operator or driver	2 (0.77)	4 (1.32)
Labourer	10 (3.83)	16 (5.28)
Other	43 (16.48)	64 (21.12)
Gaming venue employee	2 (0.77)	1 (0.33)
Ethnicity		
Australian	0 (0.00)	0 (0.00)
Indigenous Australian or Torres Strait Islander	0 (0.00)	0 (0.00)
New Zealander	0 (0.00)	0 (0.00)
Asian	24 (9.20)	19 (6.27)
Indian	0 (0.00)	5 (1.65)
Middle Eastern	3 (1.15)	2 (0.66)
European	40 (15.33)	53 (17.49)
North American	175 (67.05)	200 (66.01)
South American	7 (2.68)	6 (1.98)
African	7 (2.68)	7 (2.31)
Other (please specify)	5 (1.92)	11 (3.63)
Country		
United States of America	249 (95.40)	297 (98.02)
United Kingdom	2 (0.77)	2 (0.66)
Ireland	0 (0.00)	0 (0.00)
Australia	1 (0.38)	0 (0.00)
New Zealand	0 (0.00)	0 (0.00)
Canada	9 (3.45)	4 (1.32)

Measures

Participants provided basic demographic details (e.g., age, country), gambling frequency and expenditure, and the frequency with which they consumed pints of beer, glasses of wine, and glasses of liquor. Participants also completed the Problem Gambling Severity Index (Ferris & Wynne, 2001), and the Alcohol Use Disorders Identification Test (Saunders et al., 1993). Both the PGSI and AUDIT were scored and coded according to published guidelines. In the analyzed sample, the internal consistency of the PGSI ($\alpha = 0.93$) and AUDIT ($\alpha = 0.90$) were extremely good. Measures of internal consistency were calculated on the final analytic sample. Both the PGSI and AUDIT were scored and coded according to published guidelines.

Procedure

Two recruitment advertisements for an online survey were placed on Amazon Mechanical Turk (MTurk). The advertisement wording was identical between the two with the exception of the target activity. The specific text read as follows:

Gambling-framed study: *“This is a study about gambling and quality of life. The purpose of the study is to assess participants’ engagement in gambling activities, and if this is related to their quality of life.”*

Alcohol-framed study: *“This is a study about alcohol consumption and quality of life. The purpose of the study is to assess participants’ consumption of alcohol, and if this is related to their quality of life.”*

Data collection occurred in two waves. In the first wave, the invitation to participate in the gambling-framed study was posted on the MTurk website. In the second wave, exactly one week later, the invitation to participate in the alcohol-framed study was placed on the same website. Participants completing the first wave were assigned an MTurk ‘qualification’

which automatically excluded them seeing and participating in the second wave posting.

Within each wave, data-collection continued until the planned sample size was reached.

Participants agreeing to participate were directed to a Qualtrics survey containing the Participant Information Statement describing the respective ‘gambling’ and ‘alcohol’ framed studies in more detail, and a consent form (see preregistered protocol). Participants were then requested to complete their allocated online surveys and could do so at their own pace using either a desktop/laptop computer or smartphone. Attention and response integrity checks were included at several points (see supplementary material for examples).

Data Cleaning & Statistical analyses

All pre-processing and analyses were conducted using R (R Core Team, 2017). After removal of participants failing any one of any one of the integrity and attention checks, the following analyses were used to test hypotheses as per the preregistered protocol. Shapiro-Wilk tests indicated that responses to the PGSI and AUDIT were non-normally distributed. Accordingly, Mann-Whitney U tests were used to compare between conditions on these variables. Chi-Square tests were used to test whether the proportion of participants within each risk level of the PGSI and AUDIT differed between those in the gambling and alcohol conditions. Post-hoc Z-tests were used when Chi-Square tests were significant. Alpha was set at 0.05. Anonymized data for this study are available at <https://osf.io/d7h6e/>.

Results

Confirmatory Analyses

Figure 1 displays the PGSI and AUDIT responses for participants in the gambling and alcohol conditions. Raincloud plots (Allen et al., 2019) are used to provide a comprehensive visualization of each individual data point, their distribution, and the measures of central tendency.

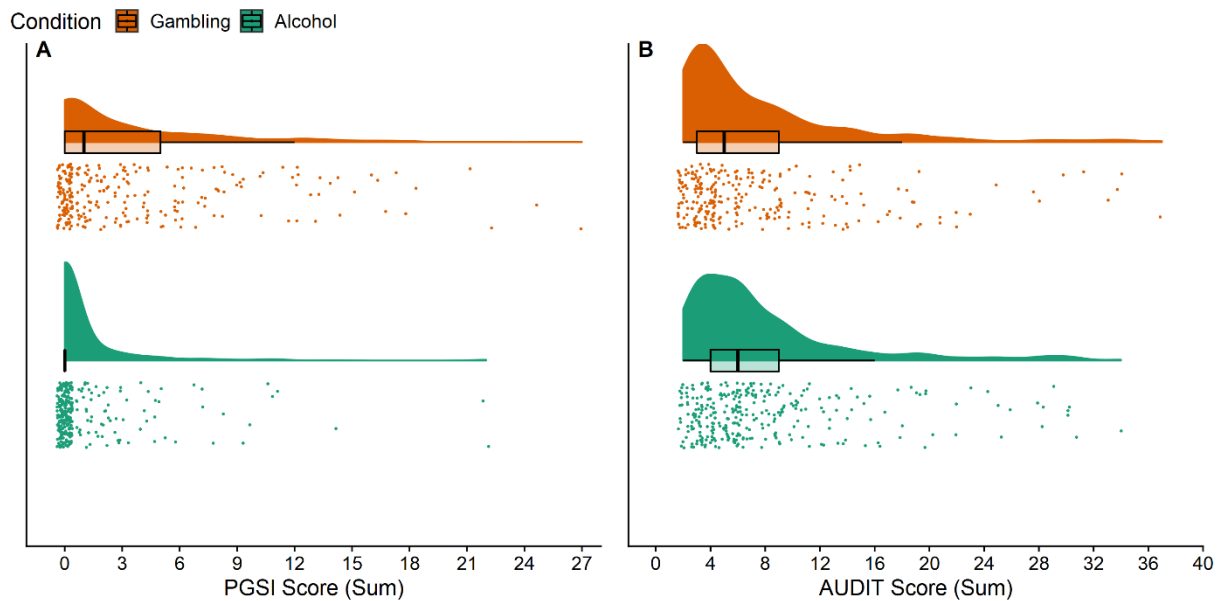


Figure 1. Raincloud plots of PGSI (A) and AUDIT (B) scores for participants in the Gambling and the Alcohol conditions. Raincloud plots simultaneously present the individual data points, the probability density, and critical summary statistics (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2019). For each condition and scale, the colored dots visualize each individual data point. The transparent curves provide a visualization of the complete distribution of the raw data. The box plots depict the measures of central tendency. To improve the visibility of individual data points, a small jitter has been applied. This does not affect the analyses of the data, only its visualization.

Consistent with our first hypothesis, participants in the gambling condition reported significantly higher PGSI scores compared to participants in the alcohol condition ($U = 55705.50$, $Z = -9.35$, $p < 0.001$, $r = 0.39$). Consistent with our second hypothesis, participants in the alcohol condition reported significantly higher AUDIT scores compared to participants in the gambling condition ($U = 35250.50$, $Z = -2.24$, $p = 0.025$, $r = 0.09$).

A similar effect was observed when examining PGSI categories as a function of condition (Table 2). Consistent with our third hypothesis, there was a significant effect of condition on PGSI category ($X^2(3, N = 564) = 88.09$, $p < 0.001$, Cramer's $V = 0.40$). A significantly higher proportion of participants in the gambling condition (16.09%) met or exceeded cut off for problem gambling compared to participants in the alcohol condition (3.30%; $p < 0.001$). A significantly higher proportion of participants in the gambling condition (24.90%) compared to the alcohol condition (8.25%) met the criteria for

classification as moderate-risk gamblers ($p < 0.001$). There was no significant difference in the proportion of low-risk gamblers across the gambling (21.07%) and alcohol conditions (13.20%; $p = 0.051$). A smaller proportion of participants in the gambling condition were non-problem gamblers (37.93%), compared to participants the alcohol condition (75.25%; $p < 0.001$).

Table 2. Distribution of Problem Gambling Classification and PGSI Scores for Gambling condition (N = 261) and Alcohol condition (N = 303) Participants

	Gambling condition (%)	Alcohol condition (%)
Non-Problem	99 (37.93)	228 (75.25)
Low Risk	55 (21.07)	40 (13.2)
Moderate Risk	65 (24.9)	25 (8.25)
Problem	42 (16.09)	10 (3.3)
PGSI Score	3.45 (4.89)	0.95 (2.67)

Note. Table only includes participants who were retained for confirmatory analyses. Mean PGSI scores are presented, with SD in parentheses.

Table 3. Distribution of Alcohol Use Disorder Classification and AUDIT Scores for Gambling condition (N = 261) and Alcohol condition (N = 303) Participants

	Gambling condition (%)	Alcohol condition (%)
Low Risk	171 (65.52)	195 (64.36)
Risky	65 (24.9)	77 (25.41)
Harmful	11 (4.21)	11 (3.63)
High Risk	14 (5.36)	20 (6.6)
AUDIT Score (SD)	7.39 (6.50)	7.85 (6.10)

Note. Table only includes participants who were retained for confirmatory analyses. Mean AUDIT scores are presented, with SD in parentheses.

Results did not support our fourth hypothesis, demonstrated by the absence of a significant effect of condition on AUDIT classification ($X^2(3, N = 564) = 0.52, p = 0.914$, Cramer's $V = 0.03$; Table 3).

Exploratory Analyses

Condition and data consistency. For the following analyses, we used the data from the 1,010 participants completing the entire survey, including the 446 participants failing any of the response integrity or attention checks. The primary purpose of these tests was to explore whether data integrity and attention varied as a function of condition. That is, were participants in the gambling condition more likely to provide inconsistent results or fail attention checks, compared to participants in the alcohol condition?

We observed that a significantly greater proportion of participants ($n = 234$) in the gambling condition had failed at least one response consistency check compared to participants in the alcohol condition ($n = 186$; $X^2(1, N = 1,010) = 8.44, p = 0.004$, Cramer's $V = 0.09$). The proportion of participants in the gambling condition that failed at least one attention check ($n = 47$) did not differ significantly from the corresponding proportion of participants in the alcohol condition ($n = 42$; $X^2(1, N = 1010) = 0.25, p = 0.620$, Cramer's $V = 0.01$).

Discussion

In this study, results provided initial evidence of a framing effect bias in MTurk studies examining problem gambling and harmful alcohol use. A significantly larger proportion of participants met criteria for a gambling problem and reported higher PGSI scores in the gambling condition compared to the alcohol condition. AUDIT scores were significantly higher in the alcohol condition; however there was no difference in terms of those meeting criteria for an alcohol problem.

Two possible explanations may underpin these results. First, participants with an interest or investment in gambling or alcohol self-select into MTurk studies based on the nature of recruitment advertisements (e.g., Williams & Volberg, 2009). Self-selection is not problematic when paired with an appropriate research question, and if the limitations are

acknowledged by researchers and understood by readers. Generalizations about prevalence rates or precise effect sizes based on self-selected samples, however, are unlikely to be accurate (Landers & Behrend, 2015). Second, participants' may have shaped their responses (consciously or unconsciously) in accordance with their beliefs about the study's purpose. There is considerable evidence demonstrating that decisions, attitudes and responses to surveys are highly influenced by how instructions and choice options are framed in respect to wording, context, and valence (positive or negative) (Kahneman & Tversky, 1979). For research on addictions, bias in self-reported data may lead to false research conclusions which has direct implications for important clinical and regulatory decisions.

Consistent with previous studies using crowdsourced samples, we observed higher rates of problem gambling and harmful alcohol use than typically reported in prevalence studies using representative sampling methods (e.g., alcohol 9.93% vs 2.2%; Glantz et al., 2020; gambling 9.22% vs 2.3%; Williams et al., 2012). Notably, our observations were still substantially below other recent crowdsourced studies, despite the use of broadly similar advertisements. For example, Kim and Hodgins (2017) used advertisements that specified drinking or gambling as the topic of investigation, observing that 66% of past 12-month drinkers reported harmful levels of alcohol use, and 34% past 12-month gamblers met criteria for problem gambling.

Notwithstanding the potential issues regarding the trustworthiness of data from MTurk (see below, and Chandler & Paolacci, 2017; Chmielewski & Kucker, 2020; Lovett et al., 2018), the high rates of behavioral and substance addiction in MTurk samples could provide a useful resource for clinical researchers to access these under-sampled populations (Shapiro et al., 2013). Recruitment of eligible participants for randomized controlled trials has historically been difficult, resulting in delayed and/or underpowered studies (Richards et al., 2014). Recent studies have demonstrated that MTurk is a viable platform to conduct

research involving online interventions, with a fast sampling speeds and high rates of participant retention at follow up (Cunningham et al., 2017, 2019). Additionally, clinical research relies mostly on treatment seeking populations which represent only a small proportion of individuals with an addictive disorder (Compton et al., 2007; Slutske, 2006). MTurk may offer a viable approach to recruiting participants who are naïve to treatment or recovering naturally. Therefore, a positive interpretation of our findings is that for clinical researchers, successful recruitment of traditionally difficult-to-access participants on MTurk may be aided by targeted and carefully worded advertising. Though, some researchers have argued those suffering the most impairment are unlikely to be reachable by survey research (see Van Rooij & Kardefelt-Winther, 2017).

The utility of MTurk as a source of under-sampled populations – using targeted advertisements or not – is only feasible if clinical cases identified are not exaggerating symptoms or outright imposters. A large proportion of participants (44.16%) in the current study failed one or more data integrity tests - substantially higher than previous crowdsourced studies. Mishra and Carleton (2017) reported 7.3% and 9.9% failed attention checks in two separate studies; however, the results of their analysis were consistent regardless of the inclusion/exclusion of these cases. Higher rates of compromised data integrity in the current study may be due to more differences in the testing of these variables compared to other MTurk studies. Participants in the gambling condition were more likely to fail the response consistency check (e.g., birth year was incompatible with their age) compared to those in the alcohol condition, suggesting gambling researchers may need to be particularly vigilant for survey response fraud.

Strengths and Limitations

This study used a tightly controlled experimental design, conducted in accordance with a preregistered protocol. We were able to utilize integrated MTurk features to ensure

that participants from the gambling condition did not have access to the alcohol condition thus avoiding contamination of results. Although our approach to recruitment ensured that the gambling and alcohol samples were independent, this method is not without limitations. First, we have may limited the total number of problem gamblers available to take part in the alcohol condition by selectively recruiting a large proportion of participants in the gambling condition who scored highly on the PGSI. The pool of participants available on MTurk may mitigate this potential problem, but we have limited evidence to support or refute this speculation. Second, we did not use random allocation to the gambling or alcohol conditions. Future studies may wish to examine whether our findings conceptually replicate when the initial study framing is neutral (e.g., “A study about recreation”) and the information statement and other study materials emphasize an addictive behavior or substance. MTurk samples were used in the current study given that this platform is the most popular crowdsource platform among researchers. Several other crowdsource websites have emerged (e.g., Figure Eight, Prolific Academic), in addition to market research panels (e.g., Qualtrics, Dynata). Results from this study cannot be generalized to other recruitment platforms. Lastly, we used a limited set of attention check items, and more up-to-date methods may mitigate data and response integrity issues (Kim & Hodgins, 2020; Mellis & Bickel, 2020).

Conclusion

MTurk is potentially a very useful resource to rapidly and inexpensively collect data from under-sampled populations. However, researchers using this recruitment strategy must be meticulous in designing the study and apply good data hygiene practices. Our results suggest that the wording of participant instructions has a significant effect on outcomes, either attracting participants with addiction-related problems or inciting response biases. Replication of the current study using alternative platforms and market research panels may help researchers select the most reliable recruitment source for paid convenience samples.

References

- Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, 4, 63. <https://doi.org/10.12688/wellcomeopenres.15191.1>
- Armstrong, A. R., & Carroll, M. (2017). *Gambling activity in Australia: Findings from wave 15 of the Household, Income and Labour Dynamics in Australia (HILDA) survey*. Australian Gambling Research Centre and the Australian Institute of Family Studies. <https://aifs.gov.au/agrc/publications/gambling-activity-australia>
- Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29(6), 2156–2160. <https://doi.org/10.1016/j.chb.2013.05.009>
- Chandler, J. J., & Paolacci, G. (2017). Lie for a dime: When most prescreening responses are honest but most study participants are impostors. *Social Psychological and Personality Science*, 8(5), 500–508. <https://doi.org/10.1177/1948550617698203>
- Chmielewski, M., & Kucker, S. C. (2020). An MTurk crisis? Shifts in data quality and the impact on study results. *Social Psychological and Personality Science*, 11(4), 464–473.
- Compton, W. M., Thomas, Y. F., Stinson, F. S., & Grant, B. F. (2007). Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: Results from the national epidemiologic survey on alcohol and related conditions. *Archives of General Psychiatry*, 64(5), 566–576.
- Cunningham, J. A., Godinho, A., & Hodgins, D. C. (2019). Pilot randomized controlled trial of an online intervention for problem gamblers. *Addictive Behaviors Reports*, 9, 100175.

- Cunningham, J. A., Godinho, A., & Kushnir, V. (2017). Can Amazon's Mechanical Turk be used to recruit participants for internet intervention trials? A pilot study involving a randomized controlled trial of a brief online intervention for hazardous alcohol use. *Internet Interventions, 10*, 12–16. <https://doi.org/10.1016/j.invent.2017.08.005>
- Dennis, S. A., Goodson, B. M., & Pearson, C. (2018). *MTurk workers' use of low-cost "virtual private servers" to circumvent screening methods: A research note*. <http://dx.doi.org/10.2139/ssrn.3233954>
- Dreyfuss, E. (2018, August 17). A bot panic hits Amazon's Mechanical Turk. *Wired*. <https://www.wired.com/story/amazon-mechanical-turk-bot-panic/>
- Ferris, J. A., & Wynne, H. J. (2001). *The Canadian problem gambling index: User manual*. Canadian Centre on Substance Abuse.
- Francalanci, C., Chiassai, S., Ferrara, G., Ferretti, F., & Mattei, R. (2011). Scale for the measurement of attitudes towards alcohol. *Alcohol and Alcoholism, 46*(2), 133–137. <https://doi.org/10.1093/alcalc/agq094>
- Gainsbury, S. M., Russell, A., & Blaszczynski, A. (2014). Are psychology university student gamblers representative of non-university students and general gamblers? A comparative analysis. *Journal of Gambling Studies, 30*(1), 11–25. <https://doi.org/10.1007/s10899-012-9334-9>
- Gainsbury, S. M., Russell, A., Hing, N., Wood, R., Lubman, D. I., & Blaszczynski, A. (2014). The prevalence and determinants of problem gambling in Australia: Assessing the impact of interactive gambling and new technologies. *Psychology of Addictive Behaviors, 28*(3), 769–779. <https://doi.org/10.1037/a0036207>
- Glantz, M. D., Bharat, C., Degenhardt, L., Sampson, N. A., Scott, K. M., Lim, C. C. W., Al-Hamzawi, A., Alonso, J., Andrade, L. H., Cardoso, G., De Girolamo, G., Gureje, O., He, Y., Hinkov, H., Karam, E. G., Karam, G., Kovess-Masfety, V., Lasebikan, V.,

- Lee, S., ... Kessler, R. C. (2020). The epidemiology of alcohol use disorders cross-nationally: Findings from the World Mental Health Surveys. *Addictive Behaviors*, 102, 106128. <https://doi.org/10.1016/j.addbeh.2019.106128>
- Grau, E., & Ortet, G. (1999). Personality traits and alcohol consumption in a sample of non-alcoholic women. *Personality and Individual Differences*, 27(6), 1057–1066. [https://doi.org/10.1016/S0191-8869\(99\)00047-1](https://doi.org/10.1016/S0191-8869(99)00047-1)
- Hays, R. D., Liu, H., & Kapteyn, A. (2015). Use of internet panels to conduct surveys. *Behavior Research Methods*, 47(3), 685–690. <https://doi.org/10.3758/s13428-015-0617-9>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). *The Weirdest People in the World?* (No. 139; Working Paper Series of the German Council for Social and Economic Data). German Council for Social and Economic Data (RatSWD). <https://ideas.repec.org/p/rsw/rswwps/rswwps139.html>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). Most people are not WEIRD. *Nature*, 466(7302), 29–29. <https://doi.org/10.1038/466029a>
- Hollingshead, S. J., Kim, H. S., Wohl, M. J. A., & Derevensky, J. L. (2016). The social casino gaming-gambling link: Motivation for playing social casino games determines whether self-reported gambling increases or decreases among disordered gamblers. *Journal of Gambling Issues*, 33, 52–67.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica (Pre-1986)*, 47(2), 263.
- Kennedy, C., Mercer, A., Keeter, S., Hatley, N., McGeeney, K., & Gimenez, A. (2016). *Evaluating online nonprobability surveys*. Pew Research Centre. <http://www.pewresearch.org/2016/05/02/evaluating-online-nonprobability-surveys/>

- Kessler, R. C., Green, J. G., Gruber, M. J., Sampson, N. A., Bromet, E., Cuitan, M., Furukawa, T. A., Gureje, O., Hinkov, H., Hu, C.-Y., Lara, C., Lee, S., Mneimneh, Z., Myer, L., Oakley-Browne, M., Posada-Villa, J., Sagar, R., Viana, M. C., & Zaslavsky, A. M. (2010). Screening for serious mental illness in the general population with the K6 screening scale: Results from the WHO World Mental Health (WMH) survey initiative. *International Journal of Methods in Psychiatric Research*, 19(S1), 4–22. <https://doi.org/10.1002/mpr.310>
- Kim, H. S., & Hodgins, D. C. (2017). Reliability and validity of data obtained from alcohol, cannabis, and gambling populations on Amazon's Mechanical Turk. *Psychology of Addictive Behaviors*, 31(1), 85–94. <https://doi.org/10.1037/adb0000219>
- Kim, H. S., & Hodgins, D. C. (2020). Are you for real? Maximizing participant eligibility on Amazon's Mechanical Turk. *Addiction*, <https://doi.org/10.1111/add.15065>
- Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology*, 8(2), 142–164. <https://doi.org/10.1017/iop.2015.13>
- Lovett, M., Bajaba, S., Lovett, M., & Simmering, M. J. (2018). Data quality from crowdsourced surveys: A mixed method inquiry into perceptions of Amazon's Mechanical Turk masters: MECHANICAL TURK DATA QUALITY. *Applied Psychology*, 67(2), 339–366. <https://doi.org/10.1111/apps.12124>
- Mellis, A. M., & Bickel, W. K. (2020). Mechanical Turk data collection in addiction research: Utility, concerns and best practices. *Addiction*. <https://doi.org/10.1111/add.15032>
- Mishra, S., & Carleton, R. N. (2017). Use of online crowdsourcing platforms for gambling research. *International Gambling Studies*, 17(1), 125–143. <https://doi.org/10.1080/14459795.2017.1284250>

- Moss, E. L., & Ranson, K. M. von. (2006). An experimental investigation of recruitment bias in eating pathology research. *International Journal of Eating Disorders*, 39(3), 256–259. <https://doi.org/10.1002/eat.20230>
- Norton, T. R., Lazev, A. B., Schnoll, R. A., & Miller, S. M. (2009). The impact of email recruitment on our understanding of college smoking. *Addictive Behaviors*, 34(6–7), 531–535. <https://doi.org/10.1016/j.addbeh.2009.03.017>
- Orford, J., Griffiths, M., Wardle, H., Sproston, K., & Erens, B. (2009). Negative public attitudes towards gambling: Findings from the 2007 British Gambling Prevalence Survey using a new attitude scale. *International Gambling Studies*, 9(1), 39–54. <https://doi.org/10.1080/14459790802652217>
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a Participant Pool. *Current Directions in Psychological Science*. <https://doi.org/10.1177/0963721414531598>
- Pardo, Y., Aguilar, R., Molinuevo, B., & Torrubia, R. (2007). Alcohol use as a behavioural sign of disinhibition: Evidence from J.A. Gray's model of personality. *Addictive Behaviors*, 32(10), 2398–2403. <https://doi.org/10.1016/j.addbeh.2007.02.010>
- Poole, J. C., Kim, H. S., Dobson, K. S., & Hodgins, D. C. (2017). Adverse childhood experiences and disordered gambling: Assessing the mediating role of emotion dysregulation. *Journal of Gambling Studies*, 33(4), 1187–1200. <https://doi.org/10.1007/s10899-017-9680-8>
- R Core Team. (2017). *R: The R Project for Statistical Computing*. <https://www.r-project.org/>
- Saunders, J. B., Aasland, O. G., Babor, T. F., De La Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol

Consumption-II. *Addiction*, 88(6), 791–804. <https://doi.org/10.1111/j.1360-0443.1993.tb02093.x>

Shapiro, D. N., Chandler, J., & Mueller, P. A. (2013). Using Mechanical Turk to study clinical populations. *Clinical Psychological Science*, 1(2), 213–220. <https://doi.org/10.1177/2167702612469015>

Skevington, S. M., Lotfy, M., O’Connell, K. A., & WHOQOL Group. (2004). The World Health Organization’s WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of Life Research*, 13(2), 299–310.

Slutske, W. S. (2006). Natural recovery and treatment-seeking in pathological gambling: Results of two U.S. national surveys. *The American Journal of Psychiatry*; Washington, 163(2), 297–302.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458. <https://doi.org/10.1126/science.7455683>

Van Rooij, A. J., & Kardefelt-Winther, D. (2017). Lost in the chaos: Flawed literature should not generate new disorders. *Journal of Behavioral Addictions*, 6(2), 128–132. <https://doi.org/10.1556/2006.6.2017.015>

Wardle, H., Great Britain, Gambling Commission, & National Centre for Social Research (Great Britain). (2011). *British gambling prevalence survey 2010*. TSO. <http://www.gamblingcommission.gov.uk/PDF/British%20Gambling%20Prevalence%20Survey%202010.pdf>

Weinstein, Y., De Lima, H. J., & van der Zee, T. (2018). Are you mind-wandering, or is your mind on task? The effect of probe framing on mind-wandering reports. *Psychonomic Bulletin & Review*, 25(2), 754–760. <https://doi.org/10.3758/s13423-017-1322-8>

- Williams, R. J., & Volberg, R. A. (2009). Impact of survey description, administration format, and exclusionary criteria on population prevalence rates of problem gambling. *International Gambling Studies*, 9(2), 101–117.
<https://doi.org/10.1080/14459790902911653>
- Williams, R. J., Volberg, R. A., & Stevens, R. M. (2012). *The population prevalence of problem gambling: Methodological influences, standardized rates, jurisdictional differences, and worldwide trends*. Ontario Problem Gambling Research Centre.
<https://www.uleth.ca/dspace/handle/10133/3068>
- Yeager, D. S., Krosnick, J. A., Chang, L., Javitz, H. S., Levendusky, M. S., Simpser, A., & Wang, R. (2011). Comparing the accuracy of RDD telephone surveys and internet surveys conducted with probability and non-probability samples. *Public Opinion Quarterly*, 75(4), 709–747. <https://doi.org/10.1093/poq/nfr020>

Supplementary Hypothesis

Additional hypotheses and analyses were included in our preregistration. Due to technical issues with the survey platform, responses to the key measure for one of these hypotheses were truncated. The following hypotheses and their corresponding analyses are included for transparency purposes only.

SH1. The gambling-framed study will be associated with more positive attitudes towards gambling compared to the alcohol-framed study.

SH2. The alcohol-framed study will be associated with more positive attitudes towards alcohol compared to the gambling-framed study.

Supplementary Method

Participants

Of the total sample, 508 completed a gambling-, and 502 an alcohol-framed questionnaire respectively. In the gambling condition, there were 276 males (54.3%) and 231 (45.5%) females and one (0.2%) non-specified gender, and 244 (48.6%) males, 254 (50.6%) females, and four (0.8%) non-specified gender in the alcohol condition. The mean age for the total sample was 37.98 years ($SD = 12.22$) with no significant age difference found between conditions (gambling: $M = 35.60$ years, $SD = 11.77$; alcohol: $M = 36.62$ years, $SD = 11.45$; Welch $t(1007.75) = -1.40$, $p = 0.161$, $d = -0.09$). The demographic details for the total sample are presented in Table S1.

The number of participants failing one or more of the response integrity checks was substantially higher than anticipated. Of the sample of 1,010 participants completing the entire questionnaire, 55.84% ($n = 564$) successfully passed all four-response integrity and three attention checks and were retained for confirmatory analyses. Of the remainder, two thirds ($n = 297$, 66.6%) failed one, a quarter ($n = 115$, 25.8%) two, and the rest ($n = 8$, 1.8%) three integrity checks. Examination of excluded participants responses to attention checks

indicated that 80.0% ($n = 357$) successfully passed all three attention checks; 17.5% ($n = 78.00$) failed one, 1.0% ($n = 4$), two, and 1.6% ($n = 7$) three attention checks, respectively.

Table S1. Demographic characteristics for Gambling condition ($N = 508$) and Alcohol condition ($N = 502$) participants.

	Gambling condition N (%)	Alcohol condition N(%)
Marital Status		
Married	210 (41.34)	186 (37.05)
Living with partner/de facto	77 (15.16)	62 (12.35)
Widowed	8 (1.58)	4 (0.80)
Divorced or separated	30 (5.91)	41 (8.17)
Never married	183 (36.02)	209 (41.63)
Education		
Year 10 or equivalent	5 (0.98)	6 (1.20)
Year 12 or equivalent	89 (17.52)	106 (21.12)
A trade/technical certificate or diploma	69 (13.58)	67 (13.35)
University or college degree	269 (52.95)	230 (45.82)
Postgraduate university qualifications	76 (14.96)	93 (18.53)
Employment Status		
Full-time	334 (65.75)	300 (59.76)
Part-time	57 (11.22)	67 (13.35)
Unemployed	29 (5.71)	42 (8.37)
Full-time student	20 (3.94)	24 (4.78)
Self-employed	42 (8.27)	40 (7.97)
Disability or other (not aged) pension	5 (0.98)	7 (1.39)
Retired (including aged pension)	16 (3.15)	12 (2.39)
Other (please specify)	5 (0.98)	10 (1.99)
Occupation		
Manager	84 (16.54)	76 (15.14)
Professional	188 (37.01)	147 (29.28)
Technician or trade worker	35 (6.89)	39 (7.77)
Community or personal service worker	21 (4.13)	23 (4.58)
Clerical or administrative	33 (6.50)	59 (11.75)
Sales	48 (9.45)	41 (8.17)
Machinery operator or driver	3 (0.59)	4 (0.80)
Labourer	22 (4.33)	22 (4.38)
Other	71 (13.98)	88 (17.53)
Gaming venue employee	3 (0.59)	3 (0.60)
Ethnicity		
Australian	2 (0.39)	1 (0.20)
Indigenous Australian or Torres Strait Islander	0 (0.00)	0 (0.00)
New Zealander	1 (0.20)	0 (0.00)
Asian	42 (8.27)	35 (6.97)
Indian	1 (0.20)	6 (1.20)
Middle Eastern	4 (0.80)	3 (0.60)
European	77 (15.16)	74 (14.74)
North American	314 (61.81)	345 (68.73)
South American	32 (6.30)	11 (2.19)
African	26 (5.12)	14 (2.79)
Other (please specify)	9 (1.77)	13 (2.59)
Country		
United States of America	476 (93.89)	486 (97.40)
United Kingdom	11 (2.17)	3 (0.60)
Ireland	0 (0.00)	0 (0.00)
Australia	4 (0.78)	1 (0.20)
New Zealand	0 (0.00)	0 (0.00)
Canada	16 (3.16)	9 (1.80)

The exclusion of 44.2% of the sample contributed to a reduction in our statistical power. The pre-registered power calculations performed using GPower 3.1 (Faul et al., 2009) indicated a sample size of 516 per condition would be required to have 95% power to detect effect sizes as small as $d = 0.22$ in independent samples t -tests. Due to the high exclusion rate, we conducted additional sensitivity analyses using GPower 3.1 that indicated that the retained sample (Gambling $N = 261$, Alcohol $N = 303$) achieved a 95% power to detect effects as small as $d = 0.30$, and 80% power to detect effects as small as $d = 0.24$ in independent samples t -tests.

Materials

Attitudes Towards Gambling Scale (ATGS; Orford et al., 2009). This 14-item scale was used to measure the extent to which individuals appraised gambling as either good or bad for society. Due to a technical error with the Qualtrics survey platform, the response options “somewhat agree” and “agree” were removed from the five-point Likert responses inadvertently truncating and biasing responses. Although the internal reliability of the ATGS in the present study was good ($\alpha = 0.87$), its factor structure is likely fatally compromised. As such, analysis of this scale is included for transparency purposes only.

Attitudes Towards Alcohol Scale (ATAS; Francalanci et al., 2011). This 14-item scale was used to measure self-reported attitudes towards alcohol. Unlike the ATGS, the ATAS does not measure attitudes towards alcohol with respect to its moral or ethical implications for society, but rather the manner in which it is used by the individual (e.g., taking advantage of free alcohol, using it to socialize). The internal consistency of the ATAS was extremely good in the present study ($\alpha = 0.93$).

Gambling Behavior. Several items were included in the survey to measure participants self-reported gambling behaviors. These included: over the previous 6 months, how many days per month they typically gambled, how many hours per day they typically

gambled for, and how much money they typically spent gambling per day. Participants were also asked to indicate whether they took part in any of the following gambling activities within the past 6 months: Keno, Bingo for money, casino table games for money, Poker for money, Electronic gaming machines, lottery or scratch tickets, horse or greyhound racing, sports betting, and other. Separate response options were available for in person engagement and online engagement. Coarse measures of engagement in in-person gambling and online gambling were calculated by taking the sum of activities participated in (e.g., Gainsbury et al., 2015).

Alcohol Consumption. Several items were also included to measure self-reported alcohol consumption. These included an item asking participants to indicate how many times in the previous week they consumed pints of beer, glasses of wine, or glasses of liquor, or other alcohol beverages (Grau & Ortet, 1999; Pardo et al., 2007). As in previous studies, we calculated an approximate total of grams of alcohol consumed per week using the formula $gr = (grade \times cc \times 0.8)/100$, where grade is the approximate alcohol content of each beverage (except for ‘other’), and cc is the typical serving size for the beverage. Note that these values are approximate only and should not be taken as being free of measurement error. The total approximate number of grams of alcohol consumed per week was calculated by summing across beer, wine, and liquor variables. Participants were also asked how frequently they consumed drinks from these categories (Never, 1-3 times per month, 1-6 times per week, Daily). An additional item was included for “other” beverages, but response to this tended to reference non-alcoholic beverages (e.g., coffee) or non-beverage intoxicants.

Kessler 6 (K6; Kessler et al., 2010). This instrument is designed to assess the presence of non-specific psychological distress experienced over the most recent four weeks. This measure was selected on the basis of its brevity, strong psychometric properties, and ability to measure severity and discriminate cases of non-specific distress from non-cases in

general-purpose health surveys. The K6 is strongly correlated with other clinical assessments of serious mental illness (Kessler et al., 2010). For each of the six K6 items covering symptoms of nervousness, hopelessness, restlessness, depression, worthlessness, and effort, response options were: ‘none of the time’ (0), ‘a little of the time’ (1), ‘some of the time’ (2), ‘most of the time’ (3) and ‘all of the time’ (4). The sum of scores on all items are calculated to provide an index of psychological distress. Internal consistency was extremely good ($\alpha = 0.92$).

WHOQOL-BREF (Skevington et al., 2004). This scale was used to measure quality of life across multiple domains and was scored according to standard guidelines. Scores for WHOQOL-BREF domains were rescaled to 0-100. Internal consistency was good (minimum $\alpha = 0.82$).

Attention Checks & Response Integrity

Three attention check items consisted of simple probe questions (e.g., “To continue, select ‘strongly agree’”) inserted into the PGSI and the AUDIT. In addition, response integrity was assessed by examining: missing responses for key variables; incomprehensible responses to open-text items; inconsistent responses to matched demographic items (e.g., indicating different ethnicities at start and end of survey); inconsistent responses to paraphrased check items from the PGSI or AUDIT; failing to successfully complete a recaptcha checkbox (confirming the absence of a bot)¹. Paraphrased PGSI and AUDIT check items were located at the end of the questionnaire and were worded to capture the same meaning as the original scale items.

Paraphrased PGSI Item: *“In the past 6 months: Have you or your household experienced any money problems because of your gambling?”*

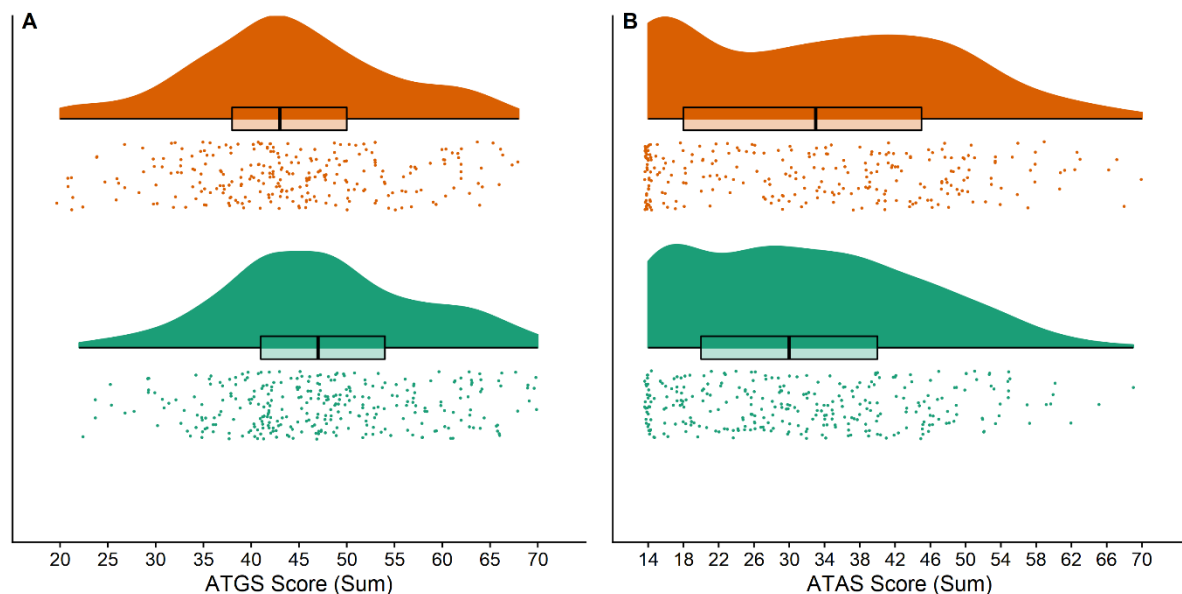
Paraphrased AUDIT Item: *“In the past 6 months how often have you: Consumed six or more alcoholic drinks in one session?”*

Data Cleaning & Statistical analyses

All pre-processing and analyses were conducted using R (R Core Team, 2017). After removal of participants failing integrity checks, the following analyses were used to test our hypotheses, as per the pre-registration protocol. Shapiro-Wilk tests were used to determine if responses to the ATGS, and ATAS violated the assumption of normality. The Shapiro-Wilk results indicated these variables were non-normally distributed. Accordingly, Mann-Whitney U tests were used to compare between conditions on ATGS, and ATAS scores. Alpha was set at 0.05.

Supplementary Results

Here, we investigated if the condition participants self-selected into was associated with differences in attitudes towards gambling or alcohol. Contrary to our supplementary hypotheses, participants in the alcohol condition reported holding significantly more positive attitudes towards gambling than did participants in the gambling condition ($U = 32469.50$, $Z = -3.67$, $p < 0.001$, $r = 0.15$). There was no significant effect of condition on attitudes towards alcohol ($U = 41305.00$, $Z = -0.92$, $p = 0.360$, $r = 0.04$).



Supplementary Figure 1: Raincloud plots of ATGS (A) and ATAS (B) scores for participants in the Gambling and the Alcohol conditions. Raincloud plots simultaneously present the individual data points, the probability density, and critical summary statistics (Allen, Poggiali, Whitaker, Marshall, & Kievit, 2019). For each condition and scale, the colored dots visualize each individual data point. The transparent curves provide a visualization of the complete distribution of the raw data. The box plots depict the measures of central tendency. To improve the visibility of individual data points, a small jitter has been applied. This does not affect the analyses of the data, only its visualization.

Exploratory Analyses

Response Quality. At the suggestion of a reviewer, we also examined whether relaxing the inclusion criteria for attention check items had any effect on our key observations. The inclusion of participants who had failed only a single attention check – and had not failed any of the response consistency checks – increased the confirmatory sample by 24 to 588. We observed the same pattern of effects of the framing manipulation on PGSI ($U = 60161.50, z = -9.27, p < 0.001, r = 0.38$) and AUDIT ($U = 37329.50, z = -2.78, p = 0.005, r = 0.11$) scores as in the pre-registered analyses.

We also explored whether the number of failed response consistency or attention checks were correlated with PGSI and AUDIT scores. Within the sample that had complete analyzable data ($n = 1010$), the number of failed response consistency scores were strongly correlated with both PGSI ($r_s = 0.52, p < 0.001$) and AUDIT ($r_s = 0.43, p < 0.001$) scores.

The number of failed attention check items were also positively correlated with PGSI ($r_s = 0.19, p < 0.001$) and AUDIT ($r_s = 0.16, p < 0.001$) scores, although only moderately so.

Gambling Behaviours and Alcohol Consumption. Participants in the gambling condition reported being significantly more involved in gambling compared to participants in the alcohol sample. As shown in Table S2, participants in the gambling condition reported more frequent gambling ($U = 56369.50, z = -9.42, p < 0.001, r = 0.40$), more hours spent gambling ($U = 55282.50, z = -8.91, p < 0.001, r = 0.38$), greater daily expenditure on gambling activities ($U = 56456.00, z = -9.37, p < 0.001, r = 0.39$), and more gambling activities engaged in person ($U = 53980.00, z = -7.97, p < 0.001, r = 0.34$) and online ($U = 46570.50, z = -5.07, p < 0.001, r = 0.21$), compared to participants the alcohol condition. There was no significant difference in the frequency or quantity of alcohol consumption for participants in the alcohol condition compared to participants in the gambling condition (Table S3).

Quality of Life. We also conducted exploratory analyses examining whether the condition participants were allocated to was associated with self-reported quality of life. As shown in Table S4, there were no significant differences in quality of life as measured by the K6 ($U = 36460.00, z = -1.60, p = 0.109, r = 0.07$), or in the physical health ($U = 41004.00, z = -0.76, p = 0.448, r = 0.03$), psychological ($U = 42846.50, z = -1.72, p = 0.086, r = 0.07$), social relationships ($U = 41708.50, z = -1.13, p = 0.258, r = 0.05$), or environment ($U = 38678.00, z = -0.45, p = 0.654, r = 0.02$) domains of the WHOQOL-BREF.

Table S2. Self-Reported Gambling Behaviours for Gambling condition (N = 261) and Alcohol condition (N = 303) Participants

	Gambling condition	Alcohol condition
Days gambled per month	3.648 (5.371)	1.297 (3.884)
Hours spent gambling	1.176 (1.380)	0.533 (1.308)
Gambling expenditure per day of gambling	93.517 (156.634)	36.109 (115.676)
Count of gambling activities (in person)	1.732 (1.792)	0.762 (1.363)
Count of gambling activities (online)	0.594 (1.083)	0.254 (0.796)

Note. Mean values are reported outside parenthesis. SD are reported inside parenthesis.

Table S3. Self-Reported Alcohol Consumption for Gambling condition (N = 261) and Alcohol condition (N = 303) Participants

	Gambling condition N (%)	Alcohol condition N (%)
Frequency of Beer drinking		
Never	110 (42.15)	115 (37.95)
1-3 times per month	96 (36.78)	116 (38.28)
1-6 times per week	49 (18.77)	70 (23.10)
Daily	6 (2.30)	2 (0.66)
$X^2 (3, N = 564) = 4.60, p = 0.203$, Cramer's $V = 0.09$		
Frequency of Wine drinking		
Never	97 (37.16)	111 (36.63)
1-3 times per month	115 (44.06)	128 (42.24)
1-6 times per week	42 (16.09)	57 (18.81)
Daily	7 (2.68)	7 (2.31)
$X^2 (3, N = 564) = 0.79, p = 0.853$, Cramer's $V = 0.04$		
Frequency of Liquor drinking		
Never	105 (40.23)	116 (38.28)
1-3 times per month	111 (42.53)	139 (45.87)
1-6 times per week	36 (13.79)	44 (14.52)
Daily	9 (3.45)	4 (1.32)
$X^2 (3, N = 564) = 3.30, p = 0.348$, Cramer's $V = 0.08$		
	<i>M (SD)</i>	<i>M (SD)</i>
Grams of alcohol consumed per week	75.276 (115.060)	79.773 (123.530)

Table S4. Self-Reported Quality of Life for Gambling condition (N = 261) and Alcohol condition (N = 303) Participants

	Gambling condition	Alcohol condition
Kessler 6	5.000 [2.000, 10.000]	6.000 [2.500, 11.000]
WHOQOL-BREF Physical Health	75.000 [64.286, 85.714]	75.000 [60.714, 85.714]
WHOQOL-BREF Psychological	66.667 [45.833, 79.167]	62.500 [41.667, 75.000]
WHOQOL-BREF Social relationships	66.667 [50.000, 75.000]	66.667 [41.667, 75.000]
WHOQOL-BREF Environment	68.750 [56.250, 81.250]	68.750 [56.250, 81.250]

Note. Median values are presented outside of brackets. Interquartile range is presented within brackets. Table only includes participants who were retained for confirmatory analyses.

Supplementary References

- Allen, M., Poggiali, D., Whitaker, K., Marshall, T. R., & Kievit, R. A. (2019). Raincloud plots: A multi-platform tool for robust data visualization. *Wellcome Open Research*, 4, 63. <https://doi.org/10.12688/wellcomeopenres.15191.1>
- Francalanci, C., Chiassai, S., Ferrara, G., Ferretti, F., & Mattei, R. (2011). Scale for the measurement of attitudes towards alcohol. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 46(2), 133–137. <https://doi.org/10.1093/alcalc/agq094>
- Gainsbury, S., Russell, A., Blaszczynski, A., & Hing, N. (2015). Greater involvement and diversity of Internet gambling as a risk factor for problem gambling. *European Journal of Public Health*, 25(4), 723–728. <https://doi.org/10.1093/eurpub/ckv006>
- Grau, E., & Ortet, G. (1999). Personality traits and alcohol consumption in a sample of non-alcoholic women. *Personality and Individual Differences*, 27(6), 1057–1066. [https://doi.org/10.1016/S0191-8869\(99\)00047-1](https://doi.org/10.1016/S0191-8869(99)00047-1)
- Kessler, R. C., Green, J. G., Gruber, M. J., Sampson, N. A., Bromet, E., Cuitan, M., Furukawa, T. A., Gureje, O., Hinkov, H., Hu, C.-Y., Lara, C., Lee, S., Mneimneh, Z., Myer, L., Oakley-Browne, M., Posada-Villa, J., Sagar, R., Viana, M. C., & Zaslavsky, A. M. (2010). Screening for serious mental illness in the general population with the K6 screening scale: Results from the WHO World Mental Health (WMH) survey initiative. *International Journal of Methods in Psychiatric Research*, 19(S1), 4–22. <https://doi.org/10.1002/mpr.310>
- Orford, J., Griffiths, M., Wardle, H., Sproston, K., & Erens, B. (2009). Negative public attitudes towards gambling: Findings from the 2007 British Gambling Prevalence Survey using a new attitude scale. *International Gambling Studies*, 9(1), 39–54. <https://doi.org/10.1080/14459790802652217>

Pardo, Y., Aguilar, R., Molinuevo, B., & Torrubia, R. (2007). Alcohol use as a behavioural sign of disinhibition: Evidence from J.A. Gray's model of personality. *Addictive Behaviors*, 32(10), 2398–2403. <https://doi.org/10.1016/j.addbeh.2007.02.010>

Behaviors, 32(10), 2398–2403. <https://doi.org/10.1016/j.addbeh.2007.02.010>

R Core Team. (2017). *R: The R Project for Statistical Computing*. <https://www.r-project.org/>

Skevington, S. M., Lotfy, M., O'Connell, K. A., & WHOQOL Group. (2004). The World Health Organization's WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of Life Research*, 13(2), 299–310.

¹ We had also pre-registered exclusion of cases with unfeasibly short or long survey response completion times. However, in our registration we neglected to operationalize the acceptable range or criteria. As such we have elected not to use this exclusion criteria.